ON LINEAR DELAYED CONTROL SYSTEMS. PART I—CONSIDERATIONS OF STABILITY

N. B. CHAKRABORTI

INSTITUTE OF RADIO PHYSICS AND ELECTRONICS, CALCUTTA UNIVERSITY

(*Received for publication August 5, 1957*)

ABSTRACT. In this paper the problem of stability of delayed control systems has been studied. The location of the roots of some characteristic equations which can be considered as the sum of a rational function and a transcendental function has been discussed. Results are derived regarding stability of systems having the characteristic equation

$$(p^3+ap^2+bp+c)e^p+rp^n = 0$$

It is shown that by the application of simple graphical methods, viz., the dual loci technique and the dual Nyquist diagram, the location of the roots can be rapidly determined and information regarding stability readily obtained.

INTRODUCTION

In a feedback control system it is required that the controlled condition should settle to the desired value after any disturbance to which the system may be subjected. It is the aim of a good control design to keep the deviation of the controlled condition to a minimum and bring it to the desired value as quickly as possible.

One of the most important factors governing the control performance is the fact that the full effect of any corrective action is not immediately felt. The delay in response assumes disturbing proportions in systems employing pneumatic, thermal and hydraulic elements. This lag is due to the finite velocity of the physical entities-pressure, or heat or mass, travelling around the loop and bears close analogy to delays in distributed parameter electrical circuits. The presence of the lag gives rise to special problems in respect of stability and time response ; the characteristic equation becomes a transcendental one and instantaneous response to a command is an impossibility.

The present author has examined in some details the problem of stability and time response of delayed control systems. In this Part I of the paper, the results of the study of stability only of certain simple systems of frequent occurrence have been discussed. The location of the roots of some characteristic equations are first considered, the principal aim being to find out the regions in the complex frequency plane where the roots with positive real parts will first appear, as the delay or the magnitude of the delayed term is varied. Analytical and graphical methods are then formulated to derive results regarding stability of certain systems.

1. *Properties of the roots of the characteristics equation*

We shall first endeavour to indicate the properties on which the boundedness of the solution of the delayed control systems depends and the regions on the complex frequency plane, where the roots of the characteristic equation with positive real parts may occur. We shall also compare the dead-time lag and distributive lag systems in respect of the location of the roots of the characteristic equation.

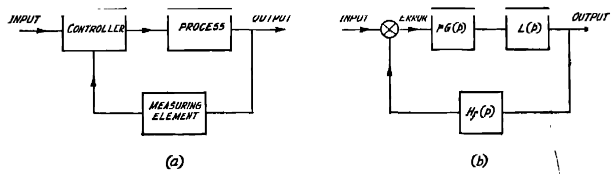


Fig 1. Block diagram of a process control system

The input-output relationship of the system of figure 1(b) is given by

$$\frac{C}{R} = \frac{G(p)L(p)r}{1+G(p)\,L(p)\,.\,rH_f(p)} = \frac{rG(p)L(p)}{1+rG(p)\,L(p)},$$

if $H_f(p) = 1$.

The characteristic roots of the system are then obtained from the zeros of

$\frac{1}{G(p)}+L(p)r.$ Putting $\frac{1}{G(p)} = H_0(p)$, we have

$$H(p) = H_0(p)+re^{-p}(\text{or } re^{-\sqrt{p}}) \qquad \ldots \quad (1a)$$

$$= H_0(p)+ rL(p) \qquad \ldots \quad (1b)$$

where $H_0(p)$ is a ratio of two polynomials, *i.e*

$$H_0(p) = \frac{\Sigma a_r p^r}{\Sigma b_r p^r} = \frac{N(p)}{D(p)} \qquad \ldots \quad (2)$$

Let n be the difference between the highest exponents of $N(p)$ and $D(p)$

Putting $p = \sigma+i\omega$ we have for a root of $H(p)$

$$A_0(\sigma, \omega)e^{i\phi(\sigma,\omega)} = -re^{-(\sigma+i\omega)}\ [\text{or} -r^{-\left[\left(\frac{\sqrt{\omega^2+\sigma^2}+\sigma}{2}\right)^{\frac{1}{2}}+i\left(\frac{\sqrt{\omega^2+\sigma^2}-\sigma}{2}\right)^{\frac{1}{2}}\right]} \ldots \quad (1c)$$

where $A_0(\sigma, \omega)$ and $\phi(\sigma, \omega)$ are the amplitudes and phase of $H_0(p)$. Let the roots of $N(p)$ and $D(p)$ be bounded by the line $\sigma = \sigma_0$ to the right. Then if $n > 0$, for $\sigma > \sigma_0$, $A_0(\sigma, \omega)$ continuously increases with σ, while $L(\sigma, \omega)$ decreases. Thus the real parts of the roots of $H(p)$ are bounded above. On the other hand, if $n < 0$, $A_0(\sigma, \omega)$ decreases with σ, and there is no upper bound of the roots of $H(p)$. We observe also that for values of σ smaller than the minimum of the real parts of the roots of $N(p)$ and $D(p)$ and $n > 0$ (which will indeed be the commonest case) both $A_0(\sigma, \omega)$ and $L(\sigma, \omega)$ increase. For large negative values of σ the amplitude equation for a dead lag system can be written as $\pi\,(\omega^2-\alpha_m) = r^2e^{-2\sigma}$, which reduces to $\omega^{2n} = r^2e^{-2\sigma}$. Hence at a root $p_k+i\omega_k$, $\omega_k >> \sigma_k$.

In a distributive lag system, for a given σ, sufficiently removed from the poles and zeros of $H_0(p)$, $A_0(\sigma, \omega)$ increases with ω, while $L(\sigma, \omega)$ decreases. And hence the possible roots are clustered round the real axis.

It is useful to note that a little removed from the rectangle containing the poles and zeros of $H_0(p)$, $\phi(\sigma\ \omega)$ is a slowly varying uniform function, so that the separation of the roots along ω of a pure lag system and along σ of a distributive lag system is approximately 2π.

Let us now consider the equation $H(p) = 0$, inside the rectangle containing all the roots of $N(p)$ and $D(p)$. We shall assume for convenience and $N(p)$ and $D(p)$ are Hurwitz. (If they are not, we have only to shift the imaginary axis to the right and our arguments and results remain valid). It is easily seen that the roots with the largest σ occur near the minima of $A_0(\sigma, \omega)$ closest to the imaginary axis. For a system with low pass properties, it is sufficient then to investigate the region near the origin in the p-plane. It is easy to observe that if $A(\sigma, \omega) > r$ for all ω, there can be no roots of $H(p)$ with non-negative real part. A root (or a pair, if complex) will occur in the neighbourhood of every zero, $\sigma+i\omega$, of $H_0(p)$-- to the right or to the left of it, depending on the phase condition.

An estimate of system performance can be obtained if the location of the first few roots in order of magnitude of the real part and the residues there at be known. It should be noted that the position of occurrence of the delay elements affects the values of the residues but not the position of the roots. It is not difficult to find the region where the real roots, if any, will occur. The complex roots are, however, less easily located. A detailed treatment of this and other related topics will be found in what follows.

2. *Roots of some transcendental equations*

We shall now consider some simple systems in respect of the location of the roots and stability.

Consider first the equation $p = c'\,e^{-p} = -c\,e^{-i\theta}e^{-p}$ where c is real. Putting $-p = Z = u+iv$, one has easily

$$v^2 = c^2e^{2u}-u^2 \qquad \ldots \quad (3a)$$

$$v \cot(v+\theta) = u \qquad \ldots \quad (3b)$$

It is easily verified that (i) the closed branch of 3(a) exists only for $c \leqslant e^{-1}$ and (ii) the whole of the open branch of 3(a) lies to the right of its intersection with the real axis. (See figure 2 where curves A_1 and A_2 correspond to 3(a) with $c = 1$

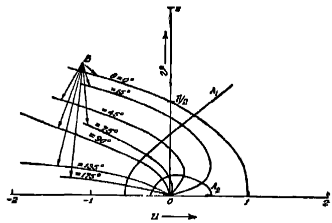


Fig. 2. Plot of (A) $v^2 = c^2 e^{2u} - u^2$
(B) $v \cot(v+\theta) = u$.

and 0.1 respectively and curves B to 3(b) for different values of θ). We note that the slope of the curve 3(a)

$$\frac{dv}{du} = \pm \frac{c^2e^{2u} - u}{\sqrt{c^2c^{2u} - u^2}}$$

is finite except at $u = ce^u$. Further, Eqns. 3(a) and 3(b) can be combined into $\log \frac{v \operatorname{cosec}(v+\theta)}{c} = v \cot(v+\theta)$. For $v > 0$, $v \cot(v+\theta)$ has a local maximum at $2v = \sin(v+\theta)$, whereafter it uniformly decreases up to $v = \pi - \theta$; while for $v < 0$, it uniformly increases, going to ∞ at $v = -\theta$. It is apparent that if $c < e^{-1}$ there will be no roots except those in the interval $-\pi+\theta < v < \pi+\theta$. Hence the condition that the roots of $p = -c_e^{-p} e^{-j\theta}$ lie to the left of $Re(p) = K$ is

$$c < e^k(v^2+k^2)^{\frac{1}{2}} \tag{4}$$

where v is a root of $v \cot(v+\theta) = -k$ and $-(\pi+\theta) < v < \pi-\theta$.

The result (4) is directly applicable to cases where the characteristic polynomial is factorisable into terms of the form $p - c_r e^{-p}$.

We now consider the equation $p_1^n = c_1 e^{-p_1\delta}$. This is reducible to $p^n = c^n e^{np}$. The curves, the intersections of which give the location of the roots are

$$v^2 = c^2e^{2u} - u^2 \tag{5a}$$

$$v \cot\left(\frac{v+2m\pi}{n}\right) = u \tag{5b}$$

where m is any integer (for $c < 0$, replaces $2m$ by $2m+1$).

Equations (5) are identical with the equations (3) if we put $\theta = \frac{2m\pi}{n}$ and the results deduced regarding the latter hold in this case also. We observe that there will be n branches of 5(b) and, in general, n roots will occur in the interval $r\pi < v < (r+1)\pi$, r being any integer, and the curve 5(b) is symmetrical about the $= 0$ axis. The root with the largest negative real part corresponds to the intersection of 5(a) with the curve 5(b) having θ or $\pi-\theta$ smallest, that is, the intersection occurring in the region $-\pi < v < \pi$. As regards boundedness, it is seen that all roots of $p^n = c^n e^{-np}$ lie to the left of $Re(p) = k$ if

$$c^2 < (v^2+k^2)\, e^{2k} \qquad \ldots \quad (5c)$$

where v is the root of $v \cot \left(v+\frac{2m\pi}{n}\right)$ in the interval $-\pi < v < \pi$. This result 5(c) is very useful when the characteristic polynomial, a part from the delay term, has a multiple root.

Next we consider the equation $p^2+2ap+b_0+r_1 p^n e^{-p}= 0$. We take first the case $n = 0$, arising in connection with a position controlled servo. On putting $p+a = z$, the equation reduces to $z^2+c+re^{-z} = 0$ where $c = b_0-a^2$ and $r_1 e^a = r$. Now putting $Z = u+iv$, we have

$$u^2+2uv \cot v-v^2+c = 0 \qquad (6a)$$

$$v^4+2v^2(u^2-c)+(u^2+c)^2-r^2e^{-2u} = 0 \qquad (6b)$$

In 6(b), in order that the real solutions of v exists, it is clearly necessary that (i) when $c > 0$, $r^2 e^{-2u} > 4cu^2$ and $u^2 < c+\sqrt{r^2e^{-2u}-4cu^2}$ and (ii) when $c < 0$, $u^2 < c+\sqrt{r^2e^{-2u}-4cu^2}$. We may recall that $c > 0$ corresponds to an under-damped two-pole structure and $c < 0$ to the overdamped case. For $c > 0$, 6(b) has cuspidal tangents at $(u^2+c)^2 = r^2e^{-2u}$, which may occur at $u > 0$ if r is large enough ($|r| < c$). If $c > |r| > 0$, 6(b) starts at the cusp, bifurcates into the upper and lower half planes. Considering the upper branch, it rises from the cusp with a positive slope, turns at $u^2+v^2 = c$ and finally goes to $u = -\infty$, $v = \infty$. The lower branch is clearly an image of the upper. The part with the largest positive u lies in the region where v lies between $\sqrt{b_0}$ and 0. If, however, $c < 0$ the curve is single valued and the part with the largest real part lies near $v = 0$. Now the curve 6(a) has two branches in each half plane. Considering the upper, if $c < 0$ or $c \leqslant 1$, the two branches start at $-1+\sqrt{1-c} = u$ and $-1-\sqrt{1-c} = u$. For $c > 1$, they start at v_0, where $\frac{\sin v_0}{v_0} = \frac{1}{\sqrt{c}}$ (clearly $0 < v_0 < \pi$) where they bifurcate and rise. We notice that at $v = (2k+1)\frac{\pi}{2}$ the two branches are at equal distances $\left(\pm\sqrt{\left((2k+1)\frac{\pi}{2}\right)^2-c}\right)$ from the $u = 0$ axis and that at $v = m\pi$ one branch is on the $u = 0$ axis and the other is at infinity, with one notable property that the slope, excepting for the part near v given by $\frac{\sin v}{v} = \frac{1}{\sqrt{c}}$, is everywhere positive. In figure 3 the values of c and r are : $c = 4$, $r = 1$; $c = 4, r = 10$; $c = -4$, $r = 10$ for the curves A_1, A_2, and A_3 respectively and the values of c for B_1 and B_2 are respectively 4 and -4.

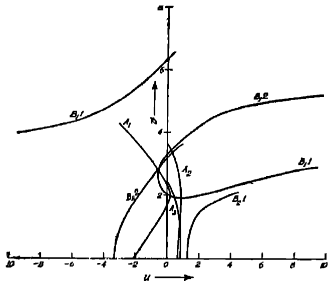


Fig. 3. Plot of (A) $\nu^4+2\nu^2(u^2-c)+(u^2+c)^2-r^2\, e^{-2u}$

(B) $\nu \cot \nu = \dfrac{v^2-u^2-c}{2u}$

From the foregoing we have the results : (i) there will, in general, be two roots of the equation $p^2+2ap+b_0+re^{-p} = 0$ in the interval $m\pi < v < (m+1)\pi$, m being any integer, (ii) the root with the largest positive real part occurs at the intersection of 6(a) and 6(b) in the region where $\pi+b_0-a^2 > |v| > 0$, if $\sqrt{b_0-a^2}$ is real; otherwise in the region $0 < v < \pi$.

Similar graphical constructions are possible for the equation $p^2+2ap+b_0+rpe^{-p} = 0$ $(n = 1)$. For this case we consider instead the mapping of the p-plane into the planes $S_1 = x_1+iy_1$ and $S_2 = x_2+iy_2$. We draw contours corresponding to u = constant and v = constant in the p-plane Since S is an analytic function of p, the transformation will be conformal and the squares in the p-plane will transform into approximate squares in the S-plane. The line $u = 0$ is the locus of S_1 at real frequencies.

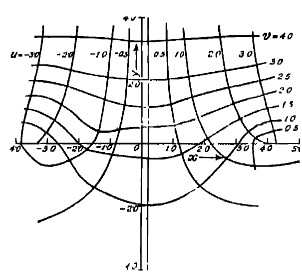


Fig 4. U-constant and v-constant contours of $\dfrac{p^2+0.2p+3}{p}$

The location of a root is found from the confluent intersections of the u-constant and v-constant curves of S_1 and S_2. It is easy to ascertain from the chart the value of r for which the largest value of the real part of the roots does not exceed, say, k. (See figure 4 which is plotted for $2a = 0.2$ and $b_0 = 3$. The plot of S_2 being very simple is not shown).

The equation $p^2+2ap+b_0+rp^2e^{-p} = 0$ ($n = 2$) can be treated on similar lines.

The study of the equation with $n > 2$ is very instructive. Considering, for example, the equation with $n = 3$, we note that the magnitude and phase condition for a root is given by

$$(b_0+2au+u^2-v^2)^2+v^2(2u+2a)^2 = (u^2+v^2)^3e^{-2u} \qquad \dots \quad (7a)$$

$$\tan^{-1}\frac{v(2u+2a)}{b_0+2au+u^2-v^2} = 3\tan^{-1}\frac{v}{u} - v + k\pi \qquad \dots \quad (7b)$$

Considering the first equality, we note that it can be satisfied for any positive u, however large, and also for small negative values ; but it cannot be satisfied for a large negative u. Noting that the curve representing the second equality extends, in general, from $u = -\infty$ to $u = \infty$ for every π interval of v, we conclude that the roots of the equation are not bounded above. This result is, in fact, true for all $n > 2$. We now consider the equation $p^n = c^{2n}e^{-2n\sqrt{p}}$. The substitution

$$\sqrt{p} = \sqrt{\sigma+i\omega} = u+iv \text{ yields}$$

$$u^2 + v^2 = c^2e^{-2u} \qquad \dots \quad (8a)$$

$$v\cot\left(\frac{v+2k\pi}{n}\right) = -u \qquad \dots \quad (8b)$$

Clearly for a large positive u, 8(a) requires $u^2+v^2 \simeq 0$, and for large negative u, $v^2 >> u^2$, hence the boundedness. The roots with positive $\sigma = R_e(p)$ correspond to the roots in the region between the lines $u = v$, $u = -v$ and the real axis.

For the case where $n = 1$, the curves we have to construct are $u^2+v^2 = c^2e^{-2u}$ and $v \cot v = -u$. These have already been drawn in another connection (see figure 2). We have here the easily verifiable result : The real part of the roots of $p = c^2e^{-2\sqrt{p}}$ will be non-positive if at the intersection of $v \cot v = -u$ and the line $u = v$,

$$c^2e^{2u} > 2u^2. \qquad \dots \quad (9)$$

3. *Stability*

We shall now consider the stability of the solution of

$$p^3 + ap^2 + bp + c + rp^ne^{-p} = 0 \qquad \dots \quad (10)$$

Coefficients a, b and c here are all real, and positive while r is real but may be either positive or negative. Such equations occur in process control systems as well as in the kinetics of circulating fuel reactors.

We refer here to Pontrajagin theorems · (I). If the polynomial $P(Z, e^Z) = P(Z, W)$ has a leading term, the real parts of the zeros of $P(Z, W)$ will be uniformly bounded above.

(II). Let $H(Z) = P(Z, e^z)$, where $P(Z, W)$ is a polynomial with a leading term. We write $H(iy) = F(y)+iG(y)$; if all the zeros of the function lie to the left of the imaginary axis then the zeros of the functions $F(y)$ and $G(y)$ are real, interlaced and

$$F(y)G'(y) - F'(y)G(y) > 0 \qquad \dots \quad (11)$$

for all values of y. Conversely, in order that the roots of the function should lie to the left of the imaginary axis, it is sufficient that at least one of the conditions given below should be satisfied :

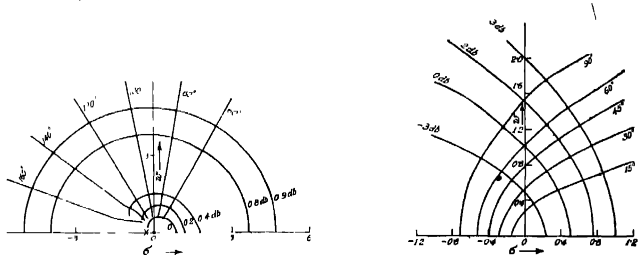


Fig. 5. (a) Equiamplitude and equiphase contours of $p^2/p+0.3$ (b) Equiamplitude and equiphase contours of $c-\sqrt{p}$

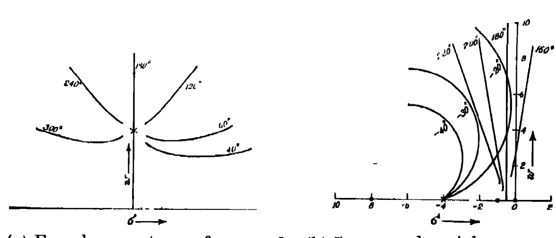


Fig. 6. (a) Equiphase contours of a pair of complex poles. (b) Superposed equiphase contours of $p(p+1)$ and $\frac{p+8}{p+4}$

(i) All the roots of the functions $F(y)$ and $G(y)$ are real and interlaced and the inequality (11) holds for at least one value of y.

(ii) All the roots of $F(y)$ are real and the inequality which now reduces to $F'(y)\ G(y) < 0$ holds for each of them.

(iii) All the roots of $G(y)$ are real the inequality (11) holds for each of them *i.e.* $G'(y)\ F(y) > 0$.

The meaning of the theorems in terms of polar plot is that the total phase $\tan^{-1}\frac{G}{F}$ should be uniformly increasing with y and that the $\mathbf{F}+\mathbf{G}$ vector will unwind itself on the y-axis as many times as it has wound itself on the infinite semi-circle to the right of the y-axis. One observes that on the semi-circle of radius $R(R\to\infty)$ the zeros of F and G occur at $(2k+1)\frac{\pi}{2} = n\theta+y$, and at $m\pi = n\theta+y$, respectively and thus each has exactly $4l\pi$ roots in an interval $2l\pi$ of y. Thus for

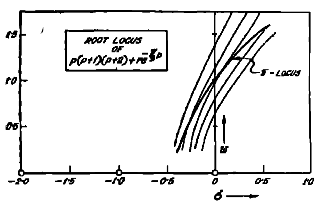
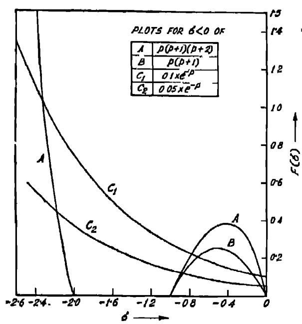


Fig. 7. (a) Root locus of $p(p+1)(p+2)+re^{-p}$



(b) Illustrates shift of real roots of $p(p+1)(p+2)+re^{-p}$ with r.

H to have no root to the right of the y-axis it is necessary that this situation be repeated on the y-axis. We now return to Eqn. (10) which can be re-written as

$$H(p) = (p^3+ap^2+bp+c)e^p+rp^n = 0 \qquad \ldots \quad (12)$$

For the presence of a leading term we require $n \leqslant 3$. Let us consider the case of $n=0$ in some detail. Writing $p = iw$ we have

$$H(iw) = F(w)+iG(w)$$
$$= (c-aw^2)\cos w+(w^3-bw)\sin w+r+i\{(c-aw^2)\sin w+(b-w^2)\,w\cos w\}$$

This can be written as

$$H(iw) = A(w)\cos(w+\phi)+r+i\ A(w)\sin(w+\phi)$$

where $A(w) = [(c-aw^2)^2+w^2(b-w^2)^2]^{\frac{1}{2}}$ and $\phi(w) = \tan^{-1}\dfrac{w(b-w^2)}{c-aw^2}$.

Then

$G'(w)\,F(w) = \{A(w)\cos(w+\phi)+r\}\{A'(w)\sin(w+\phi)+(1+\phi'(w))\,A(w)\cos(w+\phi)\}.$

We may note that the positiveness of the slope of $\phi(w)$ is ensured by the fact that the roots of $p^3+ap^2+bp+c=0$ lie in the left half of the p-plane. Since $\phi'(w)$ is always positive, the condition $F(w)\,G'(w)>0$ at the roots of G *i.e.* at $w\cot w=\dfrac{c-aw^2}{w^2-b}$, requires that

$$\cos(w+\phi)\,.\left[\cos(w+\phi)+\frac{r}{A(w)}\right]>0.$$

The inequality is best tested near the minimum of $A(w)$. Now there are two cases to consider . (a) when the roots are all real (b) when a pair is complex. If the roots are all real, it suffices to test the inequality in the region $0\leqslant w\leqslant 3\pi$ If a pair be complex, a minimum exists if the quadratic $3x^2+2(a^2-b)x+b^2-2ac$ has a real positive root. Also k defined by $w+\phi=k\pi$, should be taken as odd if $r>0$, or else even.

Now let the maximum value of $u=\dfrac{r}{A(w)}$ occur at w_0 (found from the condition that at w_0 the discriminant of the cubic $x^3+(a^2-2b)x^2+(b^2-2ac)x+c^2+\dfrac{r^2}{u^2}$ is zero). Then the condition that $H(p)$ has no root with positive real part is : At the root of $w\cot w=\dfrac{c-aw^2}{w^2-b}$ closest to the maximum of $\dfrac{r}{A(w)}$,

$$1-\frac{r}{A(\omega)}>0 \qquad \dots \quad (13)$$

We next consider Eqn. (12) for $n=1$. Thus $H(h)=(p^3+ap^2+bp+c)e+rp$. From the condition $G(w)F'(w)<0$ at the roots of $F(w)$ *i.e.* at $w\tan w=\dfrac{c-aw^2}{b-w^2}$ it follows that

$$\{A'(w)\cos(w+\phi)-A(w)\,.\,(1+\phi'(w))\sin(w+\phi)\}\{A(w)\sin(w+\phi)+rw\}<0.$$

Remembering that at a root of $F(w)$, $w+\phi=(2k+1)\dfrac{\pi}{2}$, the condition reduces to $1>\dfrac{rw}{A(w)}$ at a root of $F(w)$. Now the maximum value of $u=\dfrac{rw}{A(w)}$ renders the discriminant of the cubic $x^3+(a^2-2b)x^2+\left(b^2-2ac-\dfrac{r^2}{u^2}\right)^2+c^2$ to zero.

It can thus be readily seen that $H(p)$ will have no root with positive real part if

$$1-\frac{rw_m}{bw_m-w_m{}^3}\cos w_m>0 \qquad \dots \quad (14)$$

where w_m is the sole positive root of $w\tan w=\dfrac{c-aw^2}{b-w^2}$ in the interval $m\pi<w$

$<(m+1)\pi$ closest to the maximum of $\dfrac{rw}{A(w)}$, and provided that k in $(w+\phi) = (2k+1)\,\pi/2$ is geven when $r < 0$ or else odd.

We now give the corresponding results for the equations with $n = 2$ and 3.

The Eqn. $(p^3+ap^2+bp+c)e^p+rp^2$ will have no root with positive real part if

$$1-\frac{rw_m^2}{c-aw_m^2}\cos w_m > 0\,. \qquad \ldots \quad (15)$$

where w_m is the root of $w\cot w = \dfrac{c-aw^2}{w^2-b}$ closest to the maximum of $u = \dfrac{rw^2}{A(w)}$, and k defined by $w+\phi = k\pi$ is odd if $r < 0$ or else even.

The eqn. $(p^3+ap^2+bp+c)e^p+rp^3$ will have no root with positive real part if

$$1-\frac{r\,w_m^3}{A(w_m)} > 0\,, \qquad \ldots \quad (16)$$

where $A(w_m)$ is the magnitude of p^3+ap^2+bp+c at $p = iw_m$ and w_m is the root of $w\tan w = \dfrac{c-aw^2}{b-w^2}$ closest to the maximum of $\dfrac{rw^3}{A(w)}$.

It should be emphasised that the technique employed here is readily adapted to higher order cases. For an example, let us take the equation

$$(p^4+ap^3+bp^2+cp+d)e^p+r = 0.$$

Here we shall have two distinct cases according as the minimum value of $(w^4-bw^2+d)^2+w^2(w^2-c)^2$ occurs at $w = 0$ or not. Let it occur at w_m. We have now to find the condition for $F(w)G'(w) > 0$ at the root of $G(w)$ *i.e.* at $w\cot w = \dfrac{w^4-bw^2+d}{w^2-c}$ closest to w_m.

4. *Graphical methods*

We shall now study the location of the roots and the stability of delayed systems by graphical methods, viz., the dual loci technique and the dual Nyquist diagram.

Dual loci technique

It is known that the amplitude and phase due to poles and zeros of a network function are analogous to the potential and stream functions respectively of a set of filamentary charges of proper sign. The equipotentials and stream function contours of a single filamentary charge consist of a system of circles and radial lines. Labelling the equipotentials in logarithmic units, it is easy to draw by simple addition the resultant equi-amplitude (V) and equi-phase (ϕ) contours due to a number of poles and zeros.

We observe that the V = constant, ϕ = constant contours of re^{-p} are a set of lines parallel to the axes while those of $re^{-\sqrt{p}}$ are given by the parabolas,

$$\sqrt{\sigma^2+w^2}+\sigma = K_1 \quad \text{and} \quad \sqrt{\sigma^2+w^2}-\sigma = K_2 \qquad \text{(figure 5b)}$$

By superposing the equi-amplitude and equi-phase curves on the same graph, one can easily obtain the roots of the characteristic equation and evaluate the residues at the roots. Obviously this method facilitates prediction of system performance with adjustment and compensation.

The inverse of the overall transfer function $G_0(p)$ is given by $\frac{1}{G_0(p)} = \frac{1}{rG(p)L} + H_f(p)$. The singularities of $\frac{1}{G(p)}$ and $H_f(p)$ are first plotted on the p-plane. The equi-phase and equi-amplitude contours for $\frac{1}{G\,H_f(p)}$ and $L(p)[re^{-p}$ or $re^{-\sqrt{p}}]$ are then drawn. The next step is to plot the phase locus, *i.e.*, the locus on which the total phase equals π radians, and to insert the gain values at selected points on it.

It is useful to note some properties of the loci. Let $H_0(p)$, to restate, be $\frac{N(p)}{D(p)}$ and let n, the difference of the orders of N and D, be positive. The asymptotic amplitude and phase characteristics are the same as those of a zero of order n at the centre of charge of the system. (The root distributions of such simple systems have already been described in Sec. 2). At any particular σ one may profitably use the Bode asymptotes with proper regard, however, to the position of the singularities with respect to the σ = constant line. The phase due to a zero at a point σ, w to its right is $\tan^{-1}\frac{w-w_k}{\sigma-\sigma_k}$ and at a point to its left is $\pi - \tan^{-1}\frac{w-w_k}{\sigma-\sigma_k}$. The phase angle of $H_0(p)$ on the real axis is π radians over the segments to the right of which the number of singularities is odd. Hence if $r > 0$, the real roots of $H_0(p)+re^{-p}$ will occur only on such segments, while the roots of $H_0(p)+re^{-\sqrt{p}}$ can occur only at the points $\sigma = -m^2\pi^2$, m is any integer.

The phase loci for complex roots of delayed systems will clearly be more bent towards the right than those of non-delayed systems. Now on any π-locus the magnitude of $H_0(p)$ may increase or decrease as the real part of p increases. Confining our attention to the locus determining the location of the predominant pair of complex roots, we note that if the locus has a positive gradient and if $H_0(p)$ decreases with $Re(p)$, instability in delayed control systems sets in at a smaller value of open loop gain and the frequency of oscillation is lower.

An increase in gain, it is evident, will cause the roots to shift to regions of higher $\dfrac{H_0(p)}{L(p)}$ and the amount of shift will depend on the nature of the system, non-delayed, pure delay or distributive. In delayed systems $L(p)$ decreases uniformly with increase of $Re(p)$. Hence if gain is reduced, a root will shift to the right if $H_0(p)$ decreases with $Re(p)$ and to the left if $H_0(p)$ increases with $Re(p)$; the former movement will be less and the latter more than that occurs in non-delayed systems in similar situations. (See figure 8 depicting the location of the real roots of $p(p+1)+re^{-p}$ and $p(p+1)(p+z)+re^{-p}$ for $r = 0.1$ and $r = 0.05$).

Dual Nyquist diagram for stability

For the purposes of a dual Nyquist diagram the characteristic equation is considered as a sum of two functions $F_1(p)$ and $F_2(p)$, and the diagram consists of two polar plots of $-F_1(j\omega)$ and $F_2(j\omega)$. Instability will be revealed by the value of the total phase change over the contour formed by the imaginary axis and the infinite semi-circle in the right-half plane. This technique permits a rapid study of stability property as one, say F_2 is varied, as well as of the amplitude response.

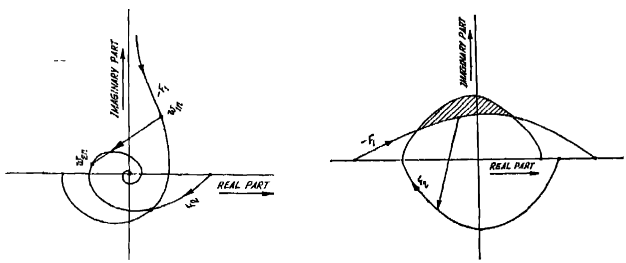


Fig. 8. A dual Nyquist diagram to illustrate
(a) Space and intersection (b) Shell.

The procedure for determining the stability of the system is stated below. On first plots the functions $-F_1(j\omega)$ and $F_2(j\omega)$. We call the points corresponding to ω_n in $-F_1$ and F_2 the terminal points ω_{1n} and ω_{2n} respectively. The region enclosed by F will be termed F-space and the region enclosed together by the arcs of F_1 and F_2 a shell. The following rules with regard to stability are easily derived:

(i) The system is stable if ω_{1n} originates inside the F_2-space and leaves it before ω_{2n} reaches the point of exit.

(ii) The system is unstable if one locus completely encloses the other.

(iii) The system is unstable if ω_{1n} enters the F_2-space and does not leave it.

(iv) The system is unstable if ω_{1n} and ω_{2n} find each other inside any shell.

We shall now consider the application of the dual Nyquist diagram in connection with the expression:

$$\frac{p^2+ap+b_0}{p^n}+rL(p)=\frac{1}{G(p)}+F(p)$$

where $L(p)$ is either e^{-p} or $e^{-\sqrt{p}}$ for certain specific cases.

1. First consider the case where $n=0$. We note that if $\left|\frac{1}{G(j\omega)}\right|>r$ for all ω there can be no intersections of the plots of $-\frac{1}{G}$ and F. Let us assume that $\left|\frac{1}{G(j\omega)}\right|<r$ for a range of values of ω. Now two distinct cases arise according as there is only one intersection or two (It is to be noted that $\left|\frac{1}{G(j\omega)}\right|$ exhibits a minimum only if $2b_0>a^2$) In the former case the system is stable if $-1/G$ leaves R_A (figure 9) before F reaches it, and in the latter the system is stable if the terminal points do not meet inside the shell $M_1\ N_1$ (figure 11).

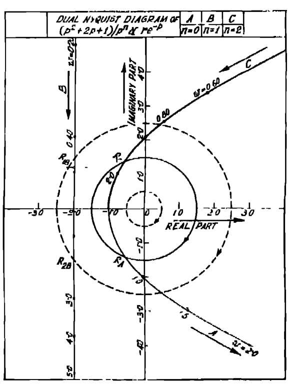


Fig. 9. Dual Nyquist diagram of $\frac{p^2+2p+1}{p^n}$ and re^{-p}

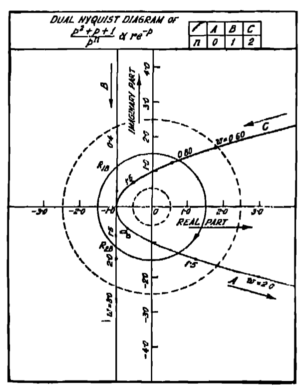


Fig 10. Dual Nyquist diagram of $\frac{p^2+p+1}{p^n}$ and re^{-p}

2. The case $n=1$. The system is stable if the terminal points do not meet inside the shell $R_{1B}-R_{2B}$ (see figures 9 and 10).

3. The case $n = 3$. The number of intersections may be one or two. In the former case it is required for stability that the terminal points be not lodged

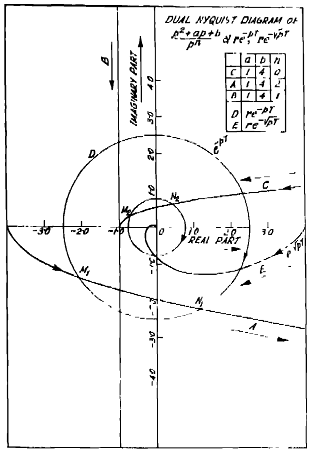


Fig. 11. Dual Nyquist diagram of $\dfrac{p^2+p+4}{p^n}$ and re^{-p}

inside $F(j\omega)$ (figure 9) and in the latter, the terminal points do not meet inside the shell M_2N_2 (figure 11).

ACKNOWLEDGMENT

The author is deeply indebted to Professor J. N Bhar for guidance and helpful suggestions and to Mr. A. K. Chowdhury and Mr. B R. Nag for valuable assistance during the progress of the work.

He is also grateful to Professor S. K. Mitra, for his encouragement and kind interest in the work.

The grant of a Senior Research Scholarship made by the Ministry of Education, Government of India, is thankfully acknowledged.

REFERENCES

Ansoff, H. I., 1949. *Jour. of Appl. Mech.* **16**, 158.

Callender, A., Hartree, D. R. and Porter, A., 1936. *Phil. Trans. Royal Soc. London*, Ser. A., **235**, p. 415-444.

Farrington, Automatic Control Systems.

Hayes, N. D., 1950. *Jour. Lond. Math. Soc.*, **25**, 226.

N. Minorsky, 1948. *Jour. of Appl. Phys.*, **19**, 332.